

Klasifikasi Lagu Berdasarkan Genre pada Format WAV

Nurmiyati Tamatjita ^{*1}, Agus Harjoko ²

¹Mahasiswa Pascasarjana UGM, ²Pascasarjana UGM; Sekip Utara Bulaksumur
Yogyakarta 55281, Telp/Fax : (0274)555133

³Jurusan Ilmu Komputer dan Elektronika, FMIPA UGM, Yogyakarta
e-mail: ^{*}tamatjita@gmail.com, ²aharjoko@ugm.ac.id

Abstrak

Dalam dunia yang berkembang pesat, media audio semakin kompleks. Karena itulah diperlukan sebuah mekanisme penentuan jenis lagu (genre) yang tepat secara efektif dan efisien. Pencarian secara manual sudah tidak efektif dan efisien lagi karena banyaknya data yang tersimpan.

Zero Crossing Rate (ZCR), Average Energy (E) dan Silent Ratio (SR) adalah 3 Feature Extraction yang digunakan untuk klasifikasi pencarian 12 genre.

Tiga dimensi adalah bentuk visualisasi pengukuran tingkat kemiripan sebuah data berdasarkan hasil klasifikasi yang diinput oleh user.

Dalam penelitian ini pengujian klasifikasi menggunakan metode 3, 6, 9 dan 12 genre melalui jarak terdekat (Euclidean Distance). Hasil pengujian yaitu menunjukkan bahwa 3 genre yaitu Balada, Blues dan Classic menunjukkan = 96,67%, 6 genre yaitu Balada, Blues, Classic, Harmony, Hip Jop dan Jazz menunjukkan = 70% dan 9 genre yaitu Balada, Blues, Classic, Harmony, Hip Hop, Jazz, Keroncong, Latin dan Pop menunjukkan = 53,33% serta 12 genre = 33,33%

Kata Kunci— Zero Crossing Rate (ZCR), Average Energy (E), Silent Ratio (SR), Euclidean Distance

Abstract

Music genre is getting complex from time to time. As the size of digital media grows along with amount of data, manual search of digital audio files according to its genre is considered impractical and inefficient; therefore a classification mechanism is needed to improve searching.

Zero Crossing Rate (ZCR), Average Energy (E) and Silent Ratio (SR) are a few of features that can be extracted from digital audio files to classify its genre. This research conducted to classify digital audio (songs) into 12 genres: Ballad, Blues, Classic, Harmony, Hip Hop, Jazz, Keroncong, Latin, Pop, Electronic, Reggae and Rock using above mentioned features, extracted from WAV audio files. Classification is performed several times using selected 3, 6, 9 and 12 genres respectively.

The result shows that classification of 3 music genres (Ballad, Blues, Classic) has the highest accuracy (96.67%), followed by 6 genres (Ballad, Blues, Classic, Harmony, Hip Hop, Jazz) with 70%, and 9 genres (Ballad, Blues, Classic, Harmony, Hip Hop, Jazz, Keroncong, Latin, Pop) with 53.33% accuracy. Classification of all 12 music genres yields the lowest accuracy of 33.33%.

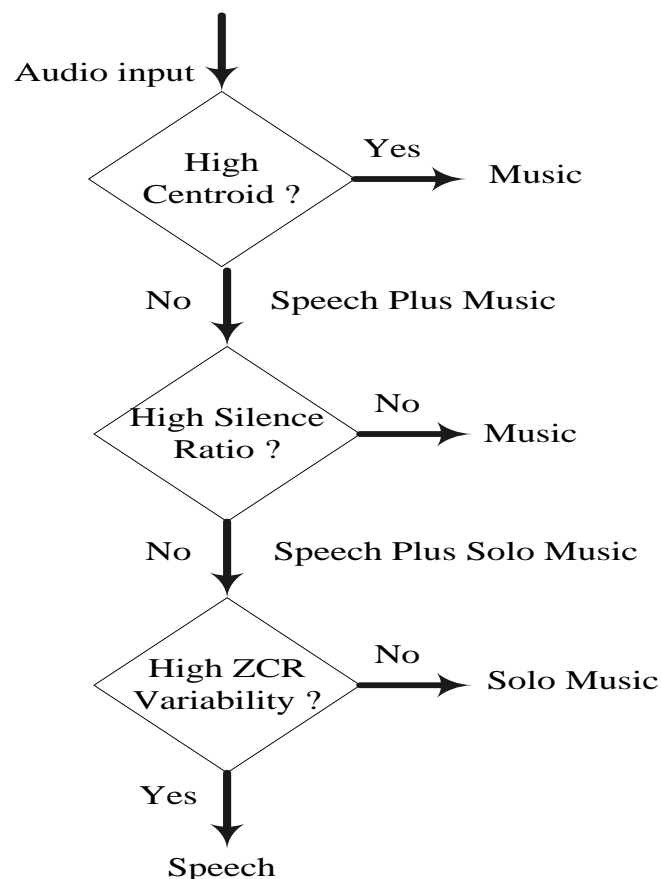
Keywords— Zero Crossing Rate (ZCR), Average Energy (E), Silent Ratio (SR), Euclidean Distance

1. PENDAHULUAN

Seiring dengan berkembangnya teknologi komunikasi berbasis digital, masyarakat membutuhkan lagu-lagu yang telah dibuat dalam bentuk digital. Musik digital menggunakan sinyal digital dalam proses reproduksi suaranya. Sebagai proses digitalisasi terhadap format rekaman musik analog, lagu atau musik digital mempunyai beraneka ragam format yang bergantung pada teknologi yang digunakan. Salah satunya adalah WAV, yang merupakan standar suara *de-facto* di Windows. Awalnya hasil *ripping* dari CD direkam dalam format ini sebelum dikonversi ke format lain. Namun sekarang tahap ini sering dilewati karena file dalam format ini biasanya tidak dikompresi dan karenanya berukuran besar.

Kendala yang timbul adalah, bagaimana mencari lagu-lagu dengan format WAV tersebut berdasarkan genre dengan tepat. Cara mula-mula ditempuh adalah dengan menduplikasikan lagu-lagu tersebut dari CD ke dalam komputer PC, yang membutuhkan banyak waktu, saat pencarian.

Klasifikasi lagu berdasarkan genre, dikembangkan untuk membantu mengatasi kendala di atas. Prinsip dasar dengan pembacaan pola melalui *amplitude* dari beberapa lagu dengan format WAV. Lagu-lagu tersebut kemudian di bagi-bagi sepanjang waktu yang dibutuhkan untuk 1 lagu. Terbentuklah pola, berdasarkan keras atau lemahnya suara dilakukan pengujian melalui beberapa sampel. Sebagai contoh adalah proses klasifikasi untuk mengelompokkan file audio yang berupa suara dan musik. Langkah-langkah klasifikasi file audio tersebut [1] ditunjukkan oleh diagram alir pada Gambar 1.



Gambar 1 Langkah-langkah Klasifikasi

Data audio yang akan diklasifikasikan memiliki beberapa perbedaan. Tabel 1 memperlihatkan contoh lain perbedaan dari *Audio Quality versus Data Rate* menurut [2].

Tabel 1 Contoh Sample Audio

Quality	Sample Rate (KHz)	Bits per Sample	Mono / Stereo	Data Rate (uncompressed) (Kb/sec)	Frequency Band (KHz)
Telephone	8	8	Mono	8	0.200-3.4
AM Radio	11.025	8	Mono	11.0	0.1-5.5
FM Radio	22.05	16	Stereo	88.2	0.02-11
CD	44.1	16	Stereo	176.4	0.005-20
DAT	48	16	Stereo	192.0	0.005-20
DVD Audio	192 (max)	24(max)	6 channels	1,200 (max)	0-96 (max)

Beberapa metode klasifikasi yang digunakan, seperti musik audio berdasarkan sistem klasifikasi simbolik. Sinyal audio adalah representasi *asymbolic* yang ditransformasikan ke harmoni menggunakan algoritma transkripsi *chord*, dengan menghitung *harmonic* kelas *pitch* profil [3]. Klasifikasi genre mengenai dua percobaan yang menggunakan mesin pembelajaran untuk mengklasifikasikan otomatis halaman web dengan genre. Eksperimen ini menyoroti pengaruh anotasi genre dan *granularity* genre dapat memiliki keakuratan untuk klasifikasi [4]. Klasifikasi genre musik melalui representasi pendengaran dari modulasi temporal dengan sebuah genre musik yang kuat, merupakan klasifikasi yang diusulkan untuk menggabungkan sifat, psiko-fisiologis didasarkan dari modulasi temporal pada rekaman musik dan kekuatan pengklasifikasi representasi berbasis jaringan. Teknik linear dimensi *subspace* pengurangan, terbukti memainkan peran penting dalam kerangka yang diteliti [5]. Klasifikasi genre musik sebagai permintaan pada pertumbuhan multimedia, pengembangan sistem pencarian informasi termasuk informasi tentang musik memiliki perhatian yang meningkat. Stasiun radio dan saluran TV musik terus melakukan pengarsipan dari jutaan kaset musik [6].

Tensor berbasis pendekatan untuk klasifikasi genre musik otomatis dengan teknik klasifikasi genre musik paling banyak menggunakan algoritma pengenalan pola untuk mengklasifikasikan fitur vektor diekstraksi dari rekaman ke dalam genre [7]. Kategorisasi genre untuk audio secara tradisional telah dilakukan secara manual. Sebuah genre musik tertentu ditandai oleh sifat statistik yang terkait dengan instrumentasi, struktur irama dan bentuk anggotanya [8].

Berdasarkan macam-macam klasifikasi tersebut, maka digunakan *Time Domain* yaitu *Zero Crossing Rate* (ZCR), *Average Energy* (E) dan *Silent Ratio* (SR) sebagai ciri untuk klasifikasi lagu, pada penelitian ini. Ciri ZCR, E dan SR ini memiliki cara pengenalan yang sederhana, yaitu dengan membagi lagu berdasarkan jenisnya, menjadi beberapa penggal lagu, kemudian menghitung nilai pada setiap segmen lagu untuk mengidentifikasi lagu tersebut. Ketiga ciri ini juga memiliki keunggulan, yaitu kemampuan menerima nilai dari sampel-sampel yang di uji dengan cara menjumlahkan.

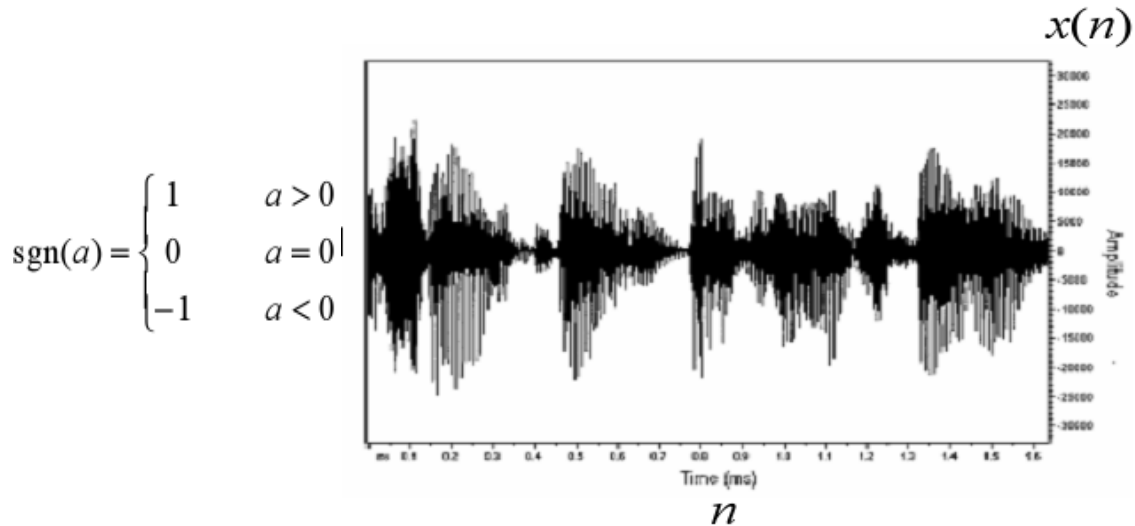
2. METODE PENELITIAN

2.1 Audio

Format berkas audio (audio file format) adalah suatu format yang digunakan dalam menyimpan berkas audio pada sistem komputer. Dalam sistem komunikasi bercirikan audio, sinyal elektrik digunakan untuk membawa unsur bunyi. Istilah ini juga biasa digunakan untuk menerangkan sistem-sistem yang berkaitan dengan proses perekaman dan transmisi yaitu sistem pengambilan/pengangkapan suara, sambungan transmisi pembawa bunyi, amplifier dan lainnya.

2.2 Format WAV

WAV adalah Waveform audio format, merupakan standar format berkas audio yang dikembangkan oleh Microsoft dan IBM. Walaupun WAV dapat menampung audio dalam bentuk terkompresi, umumnya format WAV merupakan audio yang tidak terkompres seperti Gambar 1.



Gambar 2
Sinyal Audio

2.3 Ciri Audio (Audio Features)

2.3.1 Zero Crossing Rate (ZCR)

Zero Crossing Rate (ZCR) mengindikasikan frekuensi dengan tanda pada saat berganti signal amplitudo. Jumlah dari ini mutlak dari tanda sampel yang ke-n dikurangi tanda sampel yang ke (n-1) di bagi dengan 2 kali banyaknya sampel, di mana tanda sampel yang ke n, akan bernilai 1, jika sampelnya positif. Minus 1, jika sampelnya negatif [1].

$$ZCR = \frac{\sum_{n=1}^N |\text{sgn } x(n) - \text{sgn } x(n-1)|}{2N} \quad (1)$$

$\text{Sgn } x(n)$ = Tanda dari $x(n)$, bernilai 1 jika $x(n)$ adalah positif dan bernilai -1 jika $x(n)$ adalah negatif seperti rumus (1).

N = Jumlah total sampel dalam potongan audio.

Di mana $\text{sgn}x(n)$ adalah tanda pada $x(n)$ dan akan menjadi 1, bila $x(n)$ adalah positif dan -1 bila $x(n)$ adalah negatif. Untuk melihat 1 sampel suara, diminuskan dengan sampel sebelumnya. Kalau sampel sekarang positif, nilainya 1, sedangkan negatif, nilainya -1.

2.3.2 Average Energy (*E*)

Menampilkan time-domain atau time-amplitude yang merupakan dasar dari teknik pengenalan signal, maka digunakan *Average Energy* dan *Zero Crossing Rate* sebagai rumusan. *Average Energy* mengindikasikan besar kecil suara pada signal audio [1]. Dalam kalkulasi yang simpel, dengan rumus (2).

$$E = \frac{\sum_{n=0}^{N-1} X(n)^2}{N} \quad (2)$$

E = Energi rata-rata pada potongan audio.

N = Jumlah total sampel dalam potongan audio.

X(n) = Nilai dari sampel ke n.

Di mana E adalah *Average Energy* pada satuan audio, N adalah Total nomor sampel di dalam satuan audio dan x(n) adalah nilai sampel pada sampel n.

2.3.3 Silent Ratio (*SR*)

Silent Ratio (SR) adalah jumlah dari banyaknya sampel yang dibawah nilai treshold tertentu, dibagi dengan banyaknya sampel [1] seperti rumus (3).

$$SR = \frac{\sum s}{\sum l} \quad (3)$$

s = Periode keheningan dalam potongan file audio

l = Panjang dari tiap potongan file audio

Gambar 2 menunjukkan SR [1]. Masukkan audio, apabila berpusat tinggi maka dikenal sebagai musik. Apabila tidak, maka perkataan dan musik. Apabila rasio keheningan tinggi, maka bukan musik, tetapi perkataan dan solo musik.

2.4 Klasifikasi

2.4.1 Definisi Klasifikasi

Klasifikasi adalah menggolongkan obyek pada kelasnya masing-masing [9]. Penentuan obyek pada kelas dilakukan dengan mengukur kemiripan antara obyek dengan kelasnya. Salah satu dasar kemiripan adalah menggunakan jarak terdekat (Euclidean Distance) seperti rumus (4) :

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (4)$$

p = larikan dari satu lagu yang memiliki banyak n ciri

q = larikan dari satu kelas yang memiliki banyak n ciri

3. HASIL DAN PEMBAHASAN

3.1 Klasifikasi Genre Data Audio

Proses klasifikasi yaitu membaca audio. Kemudian dilakukan pemotongan setiap 30 mili detik untuk masing-masing lagu. Hitung nilai ZCR, E dan SR untuk tiap potongan lagu tersebut, ulangi hingga potongan lagu yang terakhir. Hitung rata-rata ZCR, E dan SR. Selanjutnya pencocokan dilakukan melalui perhitungan jarak Euclidean, untuk diklasifikasikan genrenya. Ulangi proses membaca audio, sampai lagu yang terakhir.

3.2 Pengujian

Dilakukan proses uji terhadap 30 lagu untuk masing-masing kelompok dengan 4 cara, yaitu 3, 9, 6 dan 12 genre. Sehingga total lagu untuk pengujian adalah 120 lagu. Pengujian dilakukan dengan membuka file berisikan 1 lagu dengan format WAV. Kemudian dilakukan proses memotong lagu sepanjang 30 mili detik yang ke-1, ke-2 sampai durasi lagu selesai dan hitung ZCR, E dan SRnya, lalu hitung total ZCR, E, SR untuk pemotongan lagu yang dipotong ke-1, ke-2 hingga terakhir. Setelah itu, hitung rata-rata ZCR, E dan SR lagu tersebut. Lakukan pencocokan dengan jarak Euclidean, dan hasil paling terdekat sampai terjauh akan ditampilkan. Kelas genre didapatkan dari hasil jarak yang paling dekat.

3.3 Hasil Pengujian

Dilakukan pengujian terhadap metode klasifikasi untuk 3-12 genre pada lagu-lagu dengan genre yang sudah diketahui menggunakan *confusion matrix* pada Tabel 2-5. Metode klasifikasi untuk 3 Genre seperti Tabel 2 menunjukkan bahwa hasil proses klasifikasi pada pengujian menggunakan 30 lagu. Maka prosentase sebesar = 96,67% genre yang dikenali.

Tabel 2
Hasil Uji 3 Genre

Genre	A	B	C
A	9	-	1
B	-	10	-
C	-	-	10

Legenda :

A = Balada, B = Blues, C = Classic

Metode klasifikasi untuk 6 Genre seperti Tabel 3 menunjukkan bahwa hasil proses klasifikasi pada pengujian menggunakan 30 lagu. Maka prosentase sebesar = 70% genre yang dikenali.

Tabel 3
Hasil Uji 6 Genre

Genre	A	B	C	D	E	F
A	4	-	-	1	-	-
B	-	4	-	1	-	-
C	-	-	5	-	-	-
D	-	-	-	5	-	-
E	-	1	-	1	3	-
F	1	4	-	-	-	0

Legenda :

A = Balada, B = Blues, C = Classic, D = Harmony,
E = Hiphop, F = Jazz

Metode klasifikasi untuk 9 Genre seperti Tabel 4 menunjukkan bahwa hasil proses klasifikasi pada pengujian menggunakan 30 lagu. Maka prosentase sebesar = 53,33% genre yang dikenali.

Tabel 4
Hasil Uji 9 Genre

Genre	A	B	C	D	E	F	G	H	I
A	0	-	-	-	-	-	-	1	-
B	-	0	-	-	-	-	-	-	1
C	-	-	11	-	-	-	-	-	-
D	-	-	-	0	-	-	-	-	1
E	-	-	-	-	1	-	-	-	-
F	-	-	-	-	-	0	-	1	-
G	-	-	-	-	-	-	0	-	4
H	1	-	-	-	-	-	4	1	-
I	-	-	-	-	-	-	-	1	3

Legenda :

A = Balada, B = Blues, C = Classic, D = Harmony,
E = Hiphop, F = Jazz, G = Keroncong, H = Latin, I = Pop

Metode klasifikasi untuk 12 Genre seperti Tabel 5 menunjukkan bahwa hasil proses klasifikasi pada pengujian menggunakan 30 lagu. Maka prosentase sebesar = 33,33% genre yang dikenali.

Tabel 5
Hasil Uji 12 Genre

Genre	A	B	C	D	E	F	G	H	I	J	K	L
A	0	-	1	-	-	-	-	1	-	-	-	-
B	-	0	-	-	-	-	-	-	-	-	1	-
C	-	-	10	-	-	-	-	-	-	-	-	-
D	-	-	-	0	-	-	-	-	-	-	1	1
E	-	-	-	-	0	1	-	-	-	1	-	-
F	-	-	-	-	-	0	2	-	-	-	-	-
G	-	-	-	-	-	-	0	-	1	1	-	-
H	1	-	-	-	-	-	1	0	-	-	-	-
I	-	-	-	-	-	-	-	2	0	-	-	-
J	-	-	-	-	-	-	1	-	-	0	-	-
K	1	-	-	-	-	-	-	-	-	-	0	-
L	-	-	-	-	-	-	-	1	-	-	1	0

Legenda :

A = Balada, B = Blues, C = Classic, D = Harmony, E = Hiphop,
F = Jazz, G = Keroncong, H = Latin, I = Pop, J = Elektronik,
K = Reggae, L = Rock

4. KESIMPULAN

Dari pengujian yang telah dilaksanakan, didapatkan hasil yaitu :

- Dengan rata-rata tingkat akurasi pengenalan ciri 3, 6, 9 dan 12 genre. Keberhasilan pengenalan ciri oleh ZCR, E dan SR ditentukan oleh jarak antar lagu. Semakin dekat jarak antar lagu, semakin mudah rangkaian lagu tersebut dikenali.
- Penambahan lagu dapat membantu meningkatkan akurasi pengenalan.
- *Time Domain* ZCR, E dan SR kurang sesuai untuk mengenali ciri yang memiliki lebih dari 3, 6 dan 9 genre yang masuk ke daerah kelas genre lainnya. ZCR, E dan SR mengalami kebingungan untuk mengklasifikasikan genre yang memiliki properti hampir serupa untuk

lebih dari 3, 6, dan 9 genre. Kerapatan antar kelas untuk masing-masing genre terjadi selama proses ekstraksi ciri sangat berpengaruh terhadap ketepatan klasifikasi genre.

5. SARAN

Untuk memperbaiki kemampuan pengenalan ciri dengan ZCR, E dan SR, penulis mengajukan saran berikut :

Menggunakan metode klasifikasi dengan ke-3 ciri diatas, tetapi di lihat melalui frekuensi. Perlu tambahan pada proses pengenalan musik yang digunakan dalam tiap genre, bukan hanya lagu saja tetapi, minimal 2 alat instrumen untuk membangun ciri, sehingga dapat dipisahkan antara musik dan vokal.

Dapat pula dilakukan pengenalan suara penyanyi di dalam lagu, apakah solo, trio, atau pun grup dipisahkan dengan instrumen. Lagu secara utuh diklasifikasikan tetapi di dalam proses ekstraksi ciri, dipisahkan antara suara penyanyi dan instrumen, agar dapat dikenali genre dengan tepat.

Beberapa genre yang memiliki kemiripan diklasifikasi menjadi 1 kelas. Cara ekstraksi ciri dapat dilakukan pada penggalan lagu yang berada di tengah-tengah / refrein, melalui pemotongan setiap lagu yang akan di input, untuk membangun ciri.

DAFTAR PUSTAKA

- [1] Lu, Guojun, 1999, *Multimedia Database Management Systems*, Artech House Inc., London, hal. 107-115.
- [2] Li, Ze-Nian dan Drew, Mark S., 2004, *Fundamentals of Multimedia*, Pearson, Prentice Hall, Upper Saddle River, New Jersey, hal. 137.
- [3] Perez-Sancho, Carlos dan David, Rizo, 2008, Genre Classification of Music by Tonal Harmony, http://eprints.pascal-network.org/archive/00005171/01/mml08_cperez.pdf, diakses tanggal 5 Juli 2013.
- [4] Santini, Marina, 2007, *Common Criteria for Genre Classification : Annotation and Granularity*, <http://www.itri.brighton.ac.uk/~Marina.Santini/#Download>, diakses tanggal 5 Juli 2013.
- [5] Panagakis, Yannis, Kotropoulos, Constatine dan Arce, Gonzalo R., 2001, Automatic Musical Genre Classification Of Audio Signals, <http://ismir2001.ismir.net/pdf/tzanetakis.pdf>, diakses tanggal 18 Juli 2013.
- [6] Talupur, Muralidhar, Nath, Suman dan Yan, Hong, 2004, Classification Of Music Genre, <http://www.cs.cmu.edu/~yh/files/GCfA.pdf>, diakses tanggal 9 Agustus 2013.
- [7] Benetos, Emmanuil dan Kotropoulos, Constantine, 2004, Music Genre Classification Via Sparse Representations Of Auditory Temporal Modulations, http://www.ece.udel.edu/about/documents/Music_Genre_Classification_via_Sparse_Representation_of_Auditory_Temporal_Modulations_EUSIPCO2009.pdf, diakses tanggal 9 Agustus 2013.
- [8] Tzanetakis, George, Esel Georg dan Cook, Perry, 2001, Automatic Musical Genre Classification Of Audio Signals, <http://ismir2001.ismir.net/pdf/tzanetakis.pdf>, diakses tanggal 9 Agustus 2013.

- [9] Duda, Ricard O., Hart, Peter E. dan Stork, David G., 2000, *Pattern Classification*, Second Edition, John Wiley dan Sons, Inc., New York, hal. 12.